COVID-19 time series forecasting using a high-order fuzzy-neuro expert system

*

1st MANIDEEP JAKKI
Electrical and Computer Dept.
University of Waterloo
Waterloo, Canada
mjakki@uwaterloo.ca

2nd SATIPLEEN KAUR
Electrical and Computer Dept.
University of Waterloo
Waterloo, Canada
s98kaur@uwaterloo.ca

3rd YOCHITHA BATHINA
Electrical and Computer Dept.
University of Waterloo
Waterloo, Canada
ybathina@uwaterloo.ca

***Abstract*—In this paper, we present a model based on hybridization of fuzzy time series theory along with other models of RNN. In fuzzy time series, the number of intervals, lengths of intervals and order have a great impact on the forecasting accuracy. So, to create the effective lengths of intervals we are going to use the Re-Partitioning Discretization (RPD) approach. Researchers have suggested that high-order fuzzy relationships improve the forecasting accuracy of the model. Therefore, in this article, we use high-order fuzzy logic to obtain more accurate forecasting results for Covid-19 dataset. In most of the fuzzy time series models current state's (right hand side fuzzy relations) fuzzified values are used to obtain the forecasting results because the predicted values lie within the sample. For the proposed model, we are going to use advance forecasting of the time series in which previous fuzzified values (left hand side of fuzzy relations) are used to propose the model. In this paper, for comparison purpose we have performed various fuzzy time series methods which includes Chen's Model, EnsembleFTS model and ProbabilisticWeightedFTS Model. In addition to fuzzy time series model we have also done forecasting using traditional method i.e. Recurrent model in which we have trained LSTMs model and Bi-directional LSTMs model. The data set which we are going to use to evaluate the model is Covid-19 Worldwide Cases to forecast the confirmed cases, recovered cases,active cases and number of deaths around the world.**

Keywords: Machine learning, Neural networks, Fuzzy systems, Forecasting, Covid-19

I. INTRODUCTION

COVID-19 is the disease which was originated in China from Wuhan city in December 2019 and spread in the entire world. In March, 2020 it was declared as Pandemic by World Health Organization (WHO). As per latest data, the most effected countries by corona Virus are United States, Brazil, India and Russia. In this paper, we are analysing the Covid-19 confirmed cases, deaths, active cases and recovered cases of United States till date (8/13/2020).

Forecasting is the method of predicting or estimating future events on the basis of past and present data, which is typically achieved by data analysis. Advanced prediction of events such as weather, stock market and diseases are important predictive fields. In the same way, for this project we have done prediction of Covid-19 confirmed cases and deaths happened for major countries which has a larger impact due to this pandemic[1-2]. In this paper, we present a forecasting model, which is developed by hybridization of Fuzzy Time Series and Recurrent Neural Networks.

Fuzzy logic is a logic method used in the Several expert systems and artificial intelligence applications. It was first introduced by the Iranian scientist Lutfi Zadeh, at the University of California, in 1965, where he developed it as a better data processing tool[16]. Fuzzy time series is another concept for solving forecasting problems in which linguistic values are the historical data. Fuzzy time series based on Zadeh 's works[16], Song and Chissom[7], first proposed a Fuzzy Time Series forecasting model that provided a theoretical framework for modelling a special dynamic process with linguistic qualities as their findings.

Determining the lengths of the intervals of the data is very important for fuzzification of time series data sets.Generally, the lengths of intervals are kept equal in fuzzy time series models. Researchers haven't mentioned why they have used equal length of intervals[1-5]. Huarge [6] mentioned that the lengths of intervals affect the results of forecasting. So, we are using the "Re-partitioning Discretization (RPD)" approach to create the effective lengths of intervals. After generating the intervals, the next step is to fuzzify the time series data set based on the time series theory. In the past, most of the fuzzy time series models used first order fuzzy relationships to improve the performance of the forecasting model[1-4,7,10]. Researchers in their papers shows that high-order fuzzy relationships increase model performance[11-14]. So that is why in this study, we had used high-order fuzzy relationships for obtaining the forecasting results. However, to compare our forecasting results with other models we had implemented Chen's, LSTM, EnsembleFTS and ProbabilisticWeightedFTS to obtain best forecasting results.

In this paper, we are for forecasting COVID-19 cases using fuzzy time series method bot with equal intervals and unequal lengths of intervals and got better result with unequal lengths of intervals. For comparison of results we have done the forecasting by using LSTMs and Bi-directional LSTMs.After this comparison we have found out that higher order fuzzy time series fuzzy-neuro models gave us best results with minimum error.

II. Related Works

Forecasts using fuzzy time series are implemented in several areas, including forecasting of university enrollments, stock market, banking, Insurance and financial forecasts. Several methods of forecasting based on this system have been proposed in the last decades. For forecasting, most of these researchers used an interval-based FTS model to manage time series fuzzification and implemented fuzzy logic relationships that can be performed on the FTS dataset. Chen[12] developed a method for predicting fuzzy time series enrollments based on the high-order.

Many researchers had proposed different hybridization-based models to solve complex forecasting problems. Cheng et al.[29] had developed a new model of stock price forecasting based on genetic algorithm hybridization with rough set theory. Kuo et al.[30] presented hybridised particle swarm optimization with Fuzzy time series to alter length of intervals in the universe of discourse. Fuzzy time series had proven its predictive efficiency as a successful new method in predicting linguistic values .For traditional models of the time series, crisp numerical values reflect the recorded values of a particular dynamic phase. But in a fuzzy time series, linguistic values represent the recorded values of a particular dynamic process. Song and Chissom presented the first forecasting model based on the fuzzy time series[1,2] and implemented the Fuzzy Time Series model using Fuzzy Relational Equations which involves max–min composition and finally used the model to forecast the University of Alabama enrollments. A method for multi-variable fuzzy forecasting was implemented by Shyi-Ming Chen and Yu-Chuan Chang[24]. Huarng endeavoured to improve the model performance based on the determination of Length of intervals[6] and by using heuristic approaches[5]. Lee and Chou[25] predicted average university error rate enrollments lower than the Chen method[3] by more accurately defining the supports for the fuzzy numbers which represent the linguistic values of the linguistic variables. Hwang et al. [4] instead of directly using the raw numeric values he used the differences of previous data as fuzzy time series.Yu[26] developed a weighted Fuzzy Time Series model to address recurrence and weighting issues in Fuzzy Time Series prediction. Cheng et al.[27] proposed a model using fuzzy clustering technique to effectively partition the data , in order to obtain less intervals. The clustering algorithm for the K-means was implemented to partition the discourse universe in[28]. Aladag et al.[31] proposed a new method that uses feed forward neural networks to describe fuzzy relationships in high order fuzzy time series. Teoh et al.[32] proposed a fuzzy-rough hybrid forecasting model, with rules are formed via rough set algorithms. Bisht et al.[33, 35] stated fuzzy time series forecasting models based on hesitant fuzzy logical relationships to handle the problem of the aforesaid non-stochastic hesitation. An aggregation operator was introduced by Gupta and Kumar[34] to combine hesitant probabilistic fuzzy elements into fuzzy elements. These models concentrate on Fuzzy Logic Relationship modelling. Nevertheless, several works began to employ the linear model to allow forecast after high-order fuzzy time series was proposed. The above mentioned are some of the background works presented by many researchers on fuzzy time series. Considering all these references we have selected to implement higher order fuzzy time series.

III. Approach

A. Fuzzy Sets and Fuzzy Time Series

A fuzzy set is special type of set which is characterized by a membership function and are assigned to each object in the set where the value of membership ranges from zero and one, this implies that values are continuous rather than being a crisp set (which has only two values either 0 or 1). Higher the value of membership function higher is its relation with the particular set. Fuzzy time series in which it uses historical data as linguistic values, as per Zadeh's[21] work in 1965 and Song, chissom[20] work which describes briefly about linguistic values in a fuzzy sets and how these are different from traditional time series. They have defined few definitions about fuzzy logic relationships which we have inculcated in this paper.The following are few fuzzy time series definitions which have been helpful in defining fuzzy sets and relationships.

Definition1: let U be Universe of discourse where fuzzy set A can be defined as:

$$A = [\mu_A(x1)/x1 + \mu_A(x2)/x2....... + \mu_A(xi)/xi] \quad (1)$$

Definition2: If there exists a fuzzy relationship R(t 1,t), such that F(t) = F(t - 1) o R(t - 1, t), where 'o' is an arithmetic operator, then F(t) is said to be caused by F(t- 1). The relationship between F(t) and F(t 1) can be denoted by:

$$F(t-1) \longrightarrow F(t). \quad (2)$$

Definition3: If F(t - 1) = Ai and F(t) = Aj. The relationship between F(t) and F(t - 1) is referred as a fuzzy logical relationship (FLR):Ai$\longrightarrow Aj$

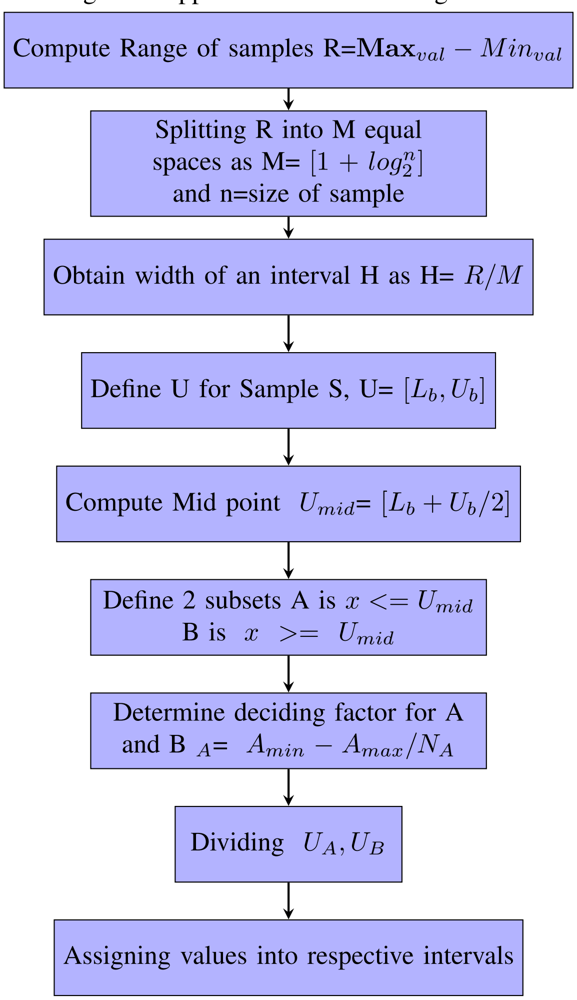
Definition4: We have implemented higher order fuzzy time series which takes the previous fuzzy data. For example, if the order is n then the relation between F(t) and previous states are as follows:

$$F(t-n), F(t-n-1).......F(t-2), F(t-1) \longrightarrow F(t) \quad (3)$$

Forecasting plays a major role in our daily life. It helps in understanding the trend for a long period of time and predict the future data by supporting in making correct decisions to reach our expectations to maximum level. We have done forecasting by applying fuzzy set theory, from the above mentioned definitions and the data is predicted by dividing into intervals and considering linguistic variables. In this paper we have proposed different time series forecasting techniques such as higher order fuzzy time series, Chen's model by using pyFTS package and other methods of RNN networks such as LSTM and compared results with each other. The following sections describes briefly about each approach.

B. Re- Partitioning Discretion(RPD)Approach

As length of intervals has the affect of RMSE and other statistical metrics of Fuzzy time series model, the choice in length interval and universal of discourse are major issues for achieving high accurate in prediction. Clustering is an technique which helps in classifying the data but it doesn't work well for few data sets. Re- Partitioning Discretion commonly known as RPD has been proposed by Singh and Borah[9] which optimized both the interval lengths and the universe of discourse. We have demonstrated the steps followed in performing RPD approach in the following flow chart:



C. Chen's Model

First forecasting was done by Song and chissom[20] applied the fuzzy times series model to propose a step-by-step procedure for forecasting the enrolments of the University of Alabama and they used the below equation for forecasting data.

$$Y(t) = Y(t-1)oR \tag{4}$$

where Y(t) is forecasting enrolment year, R is the union and 'o' is the min max operator. Later in year 1996 chen [2] proposed arithmetic operation instead of min-max composition for defuzzification which has better results and this technique has been popularly used by most of the time series forecasting for defuzzification process. In the next section we will find the model implementation for covid-19 dataset and the forecasted results compared with other proposed models.

D. Recurrent Neural Networks: LSTMs and Bi-Directional LSTMs

In real life, when a person is reading an article, want to invest in stock market or want to know about weather then he doesn't have to start each activity from the scratch. The traditional neural networks are not able to store the data and predict the next data, so to overcome this issue Recurrent Neural Networks were discovered by John Hopfields in 1982. RNNs are the networks which remembers the past and gives the output based on past experiences. The RNNs are generally used in the situation where the data is given in form of the series and the length of the series is not defined.There are different types of Recurrent neural Networks but in this we are using LSTMs and Bi-directional LSTMs.

a) LSTMs:: Long Short Term Memory Networks are the special type of Recurrent Neural Networks which are capable of learning long term connections. LSTMs were first introduced by Hochreiter and Schmidhuber in 1997 and were very useful for many problems. The LSTMs were mainly designed to overcome the long term dependency issue. To remember the information for long period of time is the default function of LSTMs which reduces the recollecting issue. Generally, all the RNNs have chain type repeating module of Neurons but in case of LSTMs the repeated chains does not only have neurons but also contains cell state and gates which allows the neural network to remember or delete the information which is required according to the desired output.In this paper we are using Vanilla LSTMs and Stacked LSTMs. a) Vanilla LSTMs: These are the models in which there is one LSTM hidden layer and output layer gives the single output. b) Stacked LSTMs: In this case we have multiple hidden layers of LSTMs one over the other and one output layer which gives us the single single output.

b) Bi-directionsl LSTMs:: These are developed as the extension of LSTMs which can improve the model for sequential input. In case where the inputs are inform of sequence then Bi-directinal LSTMs train two LSTMs model instead of one.There are times when in addition of leaning from the past to predict future we also want to look into future to fix the past so in that case Bi-directional LSTMs come into picture. Here, for this paper we have used Bi-directional LSTMs to do the comparison between the above defined methods and Bi-directional LSTMs to find out which performs better.

E. EnsembleFTS

This concept was proposed by Ahmed Mohammed[18] as an ensemble learning approach for probabilistic forecasting based on k-Nearest Neighbours, Regression Trees, Random Forests and regression methods.[17] This is a meta model which has different base models, In this model we can explore different prediction intervals by changing mode value such as extremum and quantile. In this paper we used extremum mode as it considers maximum and minimum values between the forecast.

F. ProbabilisticWeightedFTS

This is a part of higher order fuzzy time series, it is proposed for probabilistic forecasting as this consists of empirical probabilities which will be derived from the defined fuzzy sets and they also have pre-defined defuzzification rules which helps in better forecasting of future data.

G. Higher Order Fuzzy Time Series

A major disadvantage of existing fuzzy forecasting methods which are based on fuzzy time series are they uses first order time series where the forecasting results are not accurate. In Higher order fuzzy time series, the FLRs with the same LHS are gathered into groups called FLR Groups (FLRG). If F(t) is caused by F(t - 1); F(t - 2); F(t - 3);..........; F(t - p), then the corresponding high order FLR proposed by Chen[2] is

$$F(t-1); F(t-2); F(t-3); ------; F(t-n) \longrightarrow F(t) \quad (5)$$

From the above equation we can say that the weight of each F(t-1) .... F(t-n) are used to obtain F(t) forecasting at time t where n is order. In this paper, as we used Re- Partitioning Discretion approach for dividing into different interval and we have the upper and lower bounds of the intervals along with the mid points. Here we have used predefined function from package pyFTS for computing higher order fuzzy time series which was proposed by chen[2] by considering the inputs as the intervals which are obtained from the described RPD approach and order greater than 1, where the historical data is considered as covid-19 data set is used to illustrated the forecasting process. This forecasting results have much better performance than existing methods. The below image gives detailed explanation about higher order algorithm.

IV. Implementation

In this paper, we are aiming to do the forecasting of COVID-19 cases by using different techniques of Artificial Intelligence. In this we are particularly focused on forecasting by using fuzzy time series and recurrent neural networks. In this section we will explain how we have implemented all these methods to forecast the confirmed cases,deaths,recovered cases and active cases due to COVID-19.

Algorithm 1 High Order FTS Algorithm

Define the universe of discourse U;
Partition U into subintervals i.e. $U = u_1, u_2, \cdots, u_n$;
Define fuzzy sets A_i on U with the membership functions μ_{A_i};
Establish high-order FLR as:

$$F(t-1), F(t-2), F(t-3), \cdots, F(t-p) \rightarrow F(t)$$

Establish the FLRGs using above FLRs. FLRGs are determined by grouping those FLRs that have the same LHS;
For example, the following FLRs:

$$\begin{array}{l} A_{i1}, A_{i2}, \quad \cdots \quad , A_{ip} \rightarrow A_{k1} \\ A_{i1}, A_{i2}, \quad \cdots \quad , A_{ip} \rightarrow A_{k2} \\ \quad\quad\quad \vdots \\ A_{i1}, A_{i2}, \quad \cdots \quad , A_{ip} \rightarrow A_{km} \end{array}$$

produces the following FLRG:

$$A_{i1}, A_{i2}, \cdots, A_{ip} \rightarrow A_{k1}, A_{k2}, \cdots, A_{km}$$

if At time t RHS contains one or more fuzzy set in the sequence, i.e.,

$$A_{i1}, A_{i2}, \cdots, A_{ip} \rightarrow A_{k1}, A_{k2}, \cdots, A_{km}$$

then forecast at time $t+1$ is

$$\begin{aligned} FVar &= \frac{\sum_{i=1}^{k} m_{ij}}{k} - m_{i1} \\ Forecast(t+1) &= RV(t-1) + FVar \end{aligned} \quad (5)$$

where $RV(t-1)$ is the real value at time $t-1$ and m_{ij} is the midpoint of the interval related to the fuzzy set A_{ij}, i.e. the defuzzified value of A_{ij}.
end if

Fig. 1. Higher order FTS Algorithm

A. Dataset

In this paper, we are using the data which is compiled by John Hopkins University Center for Systems Sciences and Engineering (JHU CCSE) since 22 January 2020. This data is compiled from several sources which includes World Heath Organization (WHO), BNO news,DXY.cn,China CDC (CCDC), Macau Government, Hong Kong Department of Health, National Health Commission of People's Republic of China (NHC),US CDC, Taiwan CDC, Australian Government Department of Health, Government of Canada, Ministry of Health Singapore (MOH), European Center for Disease Prevention and Control (ECDC) and many more official bodies which are collecting the COVID-19 data from different countries. JHU CCSE maintains and updates the data of COVID-19 in the Data Repository of GitHub.

For our paper, we have used the data directly from the GitHub Repository so that we can get the updated data everyday. In this paper, we are using dataset of confirmed Cases and total number of deaths caused by Covid-19 pandemic.

1) Confirmed Cases: In this project we are using the confirmed cases for three countries which includes USA, Canada, Australia, India and China so that we can study the trend of number of people effected by Corona Virus in different parts of the world starting from the country from where it originated.

2) Number of Deaths: In the similar way, we are considered the number of deaths happened due to corona virus in USA, Canada, Australia, India and China to know the trend of deaths happened in these countries.

B. Data Processing

In this project, as we are using fuzzy time series approach and Recurrent Neural Networks for forecasting the cases of COVID-19. So, for both the methods the pre-processing of data is not done in the similar manner.

1) Fuzzy Time Series: We have considered the data set from john Hopkins university and the GitHub link is considered as the base URL where the data is updated on daily basis. We defined a function which extracts the confirmed cases column for all the countries. Similarly for recovered and deaths, as the data contains latitude and longitude and other columns we have dropped them and firstly we considered the confirmed cases in US country and the data is used to find the intervals using RPD approach by following the process defined in flow chart and then the data is forecasted using fuzzy logic time series. In the same way we took different countries such as India, Canada, China and Australia and found the forecasting for confirmed, deaths and recovered and active cases by following the same steps.

2) Recurrent Neural Network: LSTM's and Bi-directional LSTM: In case of Recurrent Neural Network our main motive is that the neural network is able to map the sequence of past observations and input and output so to do the same we have to do the preprocessing of the data before passing it through the model.

1) Loading the Data: First we are loading the data of confirmed cases and number of deaths directly from the data repository of GitHub created by JHU CCSE. And the extracting the data of the particular country (USA, Canada, Australia, India and China) for which we have to do the analysis.

2) Splitting the Univariant sequence: In this we are splitting the univariant sequence into multiple samples in which each sample has a specified number of time steps in out case it is 5 time steps and the output is the single timestep.

3) Now, after splitting the data we have to pass it through LSTM layer so for that we have to convert the data into three dimensional so for that we have reshaped the data.

C. Fuzzy Time Series

As defined in the above flow chart regarding RPD approach we have considered the confirmed cases of US which was used as fuzzy sets to find the intervals using defined approach after find the intervals, we get sets by describing maximum, minimum and mid-value of the interval as shown in below image:

A comparative graph is plotted as shown below between different Fuzzy time series which describes about the performance of each model with the help of the rmse, mape and U metric values which are helpful in finding the best model.

a) Chen's Model: By using Fuzzy time series using forecasting as mention in the above section we have used chen's model for forecasting, here we have used pre-defined function from pyFTS package by defining the model chen. ConventionalFTS where the partitioner parameter which is obtained by RPD approach is passed. Later the chen's model

	variables	low	high	mid_points
0	A0	-1.588671e+04	1.000000e+00	-7.942854e+03
1	A1	1.000000e+00	1.588871e+04	7.944854e+03
2	A2	1.588871e+04	3.177642e+04	2.383256e+04
3	A3	3.177642e+04	4.766413e+04	3.972027e+04
4	A4	4.766413e+04	6.355184e+04	5.560798e+04
...	...	...	...	...
96	A96	3.478236e+06	3.536259e+06	3.507248e+06
97	A97	3.536259e+06	3.594283e+06	3.565271e+06
98	A98	3.594283e+06	3.652306e+06	3.623294e+06
99	A99	3.652306e+06	3.710329e+06	3.681317e+06
100	A100	3.710329e+06	3.768352e+06	3.739341e+06

Fig. 2. Higher order FTS Algorithm

is fitted to covid confirmed dataset, a proper model with hyper-parameter tuning is will produce accurate results. In this approach we have forecasted the true values of confirmed cases along with the predicted values from this method, from the below graph we can depict that chen's model had good forecasting at lower values but there are ambiguities at forecasting higher values as shown in the below figure.

b) EnsembleFTS Model: The forecasting procedure consists of evaluating the models and combining their outputs into a continuous probability distribution by considering Kernel density estimation. In this method we have defined the model ensemble.AllMethodEnsembleFTS from pyFTS package and the partitioner parameter is passed as similar to the above defined model. In the below graph we can say that ensemble method fits well for the given data but there are fluctuations in the predicted graph and the rmse values are high when compared to other models

c) ProbabilisticWeightedFTS Model: In this method the interval are calculated by considering the fuzzy sets and its weights of the probability. In this method we defined a predefined model from pyFTS package and we have defined partitioner and heuristic as the hyper parameters for better forecasting. In the below graph we can say that this method fits well for the given data but due to considering the weights of the probabilities there is change in the forecasting graph for higher values.

d) Higher Order Fuzzy Time Series: Based on the above mention definition of high-order fuzzy time series, order from 1 to 7 has been established in the paper. For example, the fuzzified values for 4th order for the date 12/August/2020 are previous 4 values from 11th august to 8th august. Here to find the 4th order fuzzy time series let's consider the 4 values as A0, A1, A2 and A3, Thus the 4th order FLR is represented

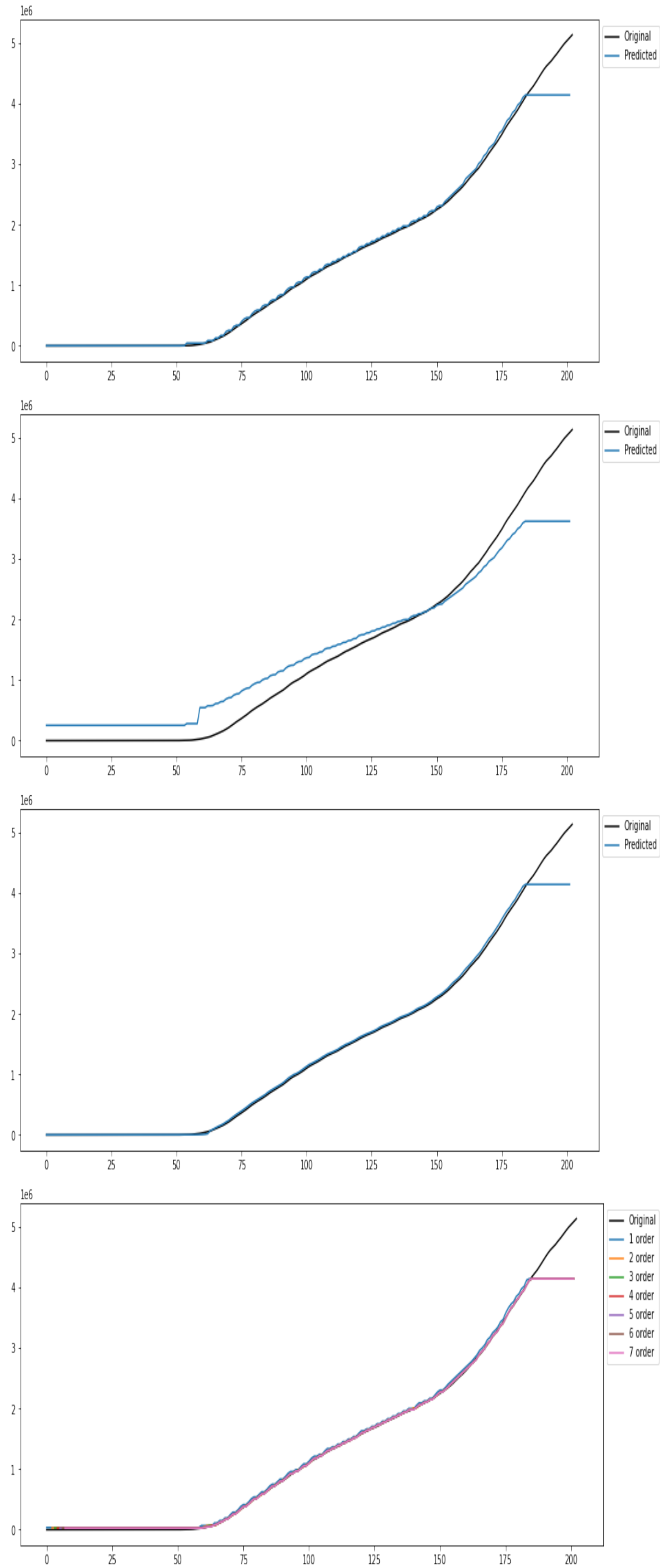


Fig. 3. (a) Chen's Model for Confirmed cases in US (b) EnsembleFTS Model for Confirmed cases in US (c) ProbabilisticWeightedFTS Model for Confirmed cases in US (d)Higher Order FTS Model for Confirmed cases in US

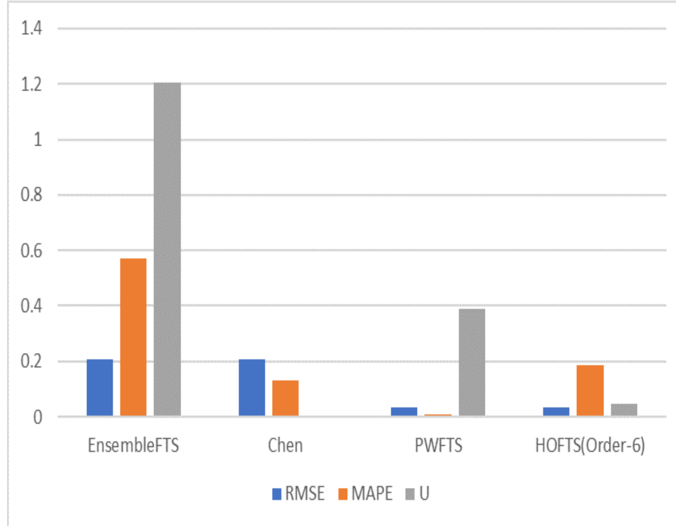


Fig. 4. Higher Order FTS Model for Confirmed cases in US

as follows:

$$A0, A1, A2, A3 \longrightarrow A4 \tag{6}$$

Here in this paper we have implemented orders from 1 to 7 by using package from pyFTS as hofts.HighOrderFTS and in this model we have parameters such as order and partitioner which are defined in the previous section. By fitting the defined model and the data is forecasted using predict function. In order to obtain different Measurements such as RMSE commonly know as Root Mean Square error which is used to calculate the difference between actual and predicted values and similarly MAPE(Mean absolute percentage error) and Theil's U statistic[24] are calculated for each order such that we can choose model which gives best forecasting results. In the below graph as the data lengths are large the graphs are overlapped for all the order a keen difference can be observed in metrics values and can easily differentiate that which order suits best for the provided dataset.

D. Recurrent Neural Network

Recurrent Neural Networks (RNN) is the subset of Artificial Neural Networks in which each connection are nodes to form a directional graph in the form of sequence. A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence which allows it to exhibit temporal dynamic behavior. Derived from feed forward neural networks, RNNs can use their internal state (memory) to process variable length sequences of inputs

a) Vanilla LSTMs: Vanilla LSTMs are the models in which there is one hidden layer of LSTMs. For this paper we have used univariant sequence. First we have loaded the Covid-19 data of confirmed cases from the data repository of JHU CCSE. Then, we have splitted the data into training and test data in 80% and 20% respectively. Here, we are taking the COVID-19 confirmed cases and it is univariant sequence. Now, we will split the train data and test data into time steps for which we have made a function in which we are passing the data set and splitting into time steps. For this we are taking size of time step to be 5 which means that the input sequence will be of 5 elements with single output. Now, we are converting the data into three dimension because we can pass only three dimensional data into LSTM layer. Here we have take one hidden layer with one feature as we are having univariant sequence. After that we are training the model using training data, Then we are finding out the mean square error. After that we are predicting the data on the test set. The forecasting cure for test data in figure-5 (a).

b) Stacked LSTMs: Stacked LSTMs are the models in which there are more than one hidden layer. In similar way as we did in Vanilla LSTM we are taking Covid-19 confirmed cases as univariant dataset. Again, we have first splitted the data into training and test data in 80% and 20% respectively. After that we will split the train and test data into time steps and in this also we are considering it to be 5 time steps. So, the input sequence will have 5 elements and output will be single value. After that again we will convert the data into three dimension because to pass the data through LSTM we need the data to be reshaped to three dimensions. Then we will train the stacked LSTM model and forecast the test data using the same. Then we have calculated the mean square error. So, after that we found that bi-direction LSTMs model performs better than traditional LSTMs model which can also be seen from the figure-5 (b).

c) Bi-directional LSTMs: In Bi-directional LSTMs the LSTM layer which we use in the model are bi-directional in nature i.e. rather than one LSTM model we are training two of them one from left to right and other one from right to left. Similar to Vanilla and stacked LSTMs here also we are considering the univariant sequence of Covid-19 confirmed cases. Here also we are first splitting the data into train and test in 80% and 20% respectively. After that we will split the train and test data into time steps and in this also we are considering it to be 5 time steps. So, the input sequence will have 5 elements and output will be single value. After that again we will convert the data into three dimension because to pass the data through Bi-LSTM we need the data to be reshaped to three dimensions. Then we will train the bi-directional LSTM model and forecast the test data using the same. Then we have calculated the mean square error. So here we found out that results by stacked LSTMs were better than vanilla LSTMs as in this we got less error than Vanilla LSTMs which can also be seen from the figure-5 (c).

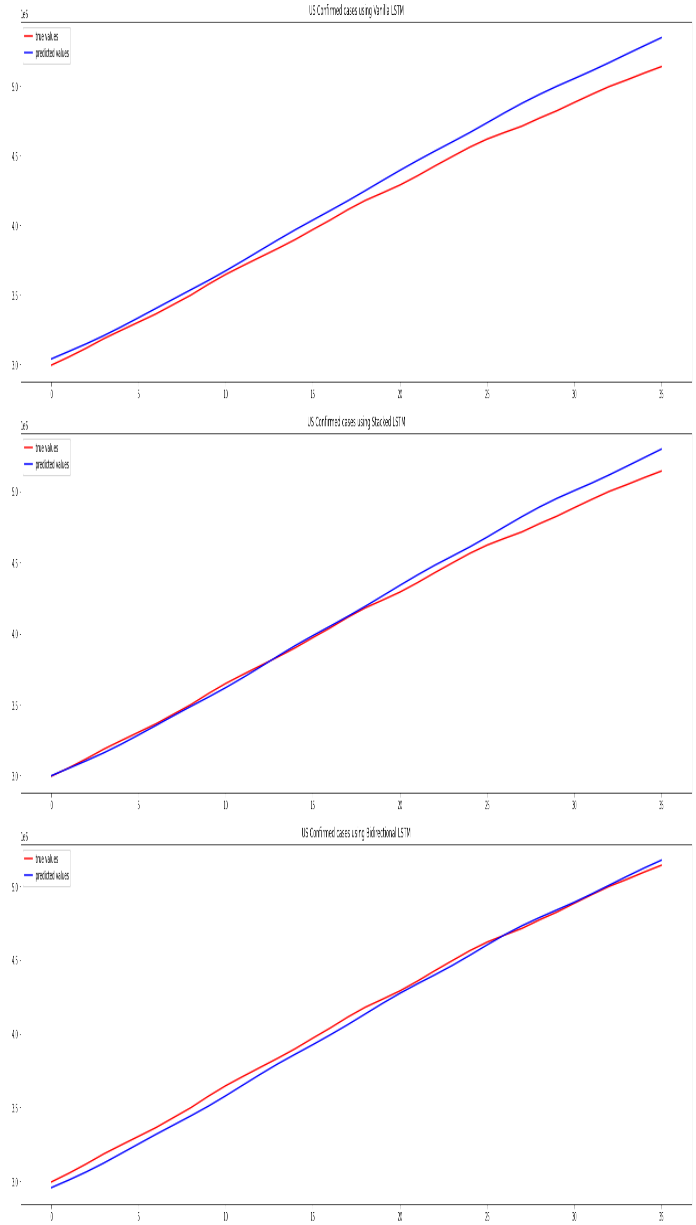


Fig. 5. (a) Vanilla LSTMs for Confirmed cases in US (b) Stacked LSTMs for Confirmed cases in US (c) Bi-directional LSTMs for Confirmed cases in US (d)Higher Order FTS Model for Confirmed cases in US

V. Results and Analysis

In this paper, we have performed higher order fuzzy time series method for COVID-19 confirmed cases, active cases, deaths and recovered cases. For comparison we have also tried with other fuzzy time series methods which are EnsembleFTS model, ProbablisticWeightedFTS model and Chen's model and done the forecasting of only confirmed cases. In this paper, we have also performed traditional methods for forecasting i.e. forecasting the confirmed cases of COVID-19 using Recurrent Neural Network for comparison with fuzzy time series techniques.

TABLE I
HIGHER ORDER FTS FOR USA

Higher Order FTS for USA												
	Confirmed Cases			Recovered Cases			Deaths			Active		
Order	**RMSE**	**MAPE**	**U**	**RMSE**	**MAPE**	**U**	**RMSE**	**MAPE**	**U**	**RMSE**	**MAPE**	**U**
1	0.2123	0.9562	0.0834	0.5232	0.9434	0.0732	0.7645	0.01	0.0623	1.00596107	0.1963042	0.0932
2	0.3343	0.9698	0.0808	0.9286	0.7662	0.0778	0.2764	0.02	0.0123	1.3261	0.5354	0.0853
3	0.4754	0.7343	0.0674	0.8544	0.3433	0.0568	0.4764	0.03	0.3223	1.4564	0.1042	0.0461
4	0.3903	0.4397	0.0506	0.5644	0.8161	0.0271	0.4482	0.01	0.0524	1.3433	1.2147	0.0322
5	0.5604	0.4382	0.0559	0.6845	0.5498	0.0453	0.6614	0.009	0.0942	1.6887	0.9947	0.0234
6	0.1217	0.3228	0.0306	0.4656446	0.8467	0.0731	0.72845	0.08	0.0492	1.5733	0.8929	0.0857
7	0.7823	0.7745	0.4536	0.486841	0.5453	0.0675	0.1336	0.10	0.0562	1.6438	0.9109	0.043

TABLE II
FORECASTING OF COVID-19 CONFIRMED CASES

Confirmed Cases Prediction for last 25 days(US)		
Dates	**True value**	**Predicted value**
7/18/20	3711413	4034645
7/19/20	3773260	4096492
7/20/20	3834677	4157909
7/21/20	3899211	4222443
7/22/20	3970121	4293353
7/23/20	4038816	4362048
7/24/20	4112531	4435763
7/25/20	4178970	4502202
7/26/20	4233923	4557155
7/27/20	4290337	3967105
7/28/20	4356206	4679438
7/29/20	4426982	4750214
7/30/20	4495015	4171183
7/31/20	4562107	4885339
8/01/20	4620592	4943824
8/02/20	4668172	4991404
8/03/20	4713540	4168996
8/04/20	4771080	5094312
8/05/20	4823890	5147122
8/06/20	4883582	5206814
8/07/20	4941755	5264987
8/08/20	4997929	5321161
8/09/20	5044864	5368096
8/10/20	5094400	5417632
8/11/20	5141208	5418976

TABLE III
COMPARISON OF FUZZY TIME SERIES MODEL

Confirmed Cases of US				
Metrics	**Chen**	**EnsembleFTS**	**ProbabilisticWeightedFTS**	**Higher OrderFTS(Order-6)**
RMSE	0.1971	0.8385	0.4347	0.1819
MAPE	0.3004	0.6998	0.8218	0.5408
U	0.3852	0.3122	0.5005	0.2039

The performance of all the models is evaluated on the basis of mean absolute percentage error of observed and predicted values, root mean square error and Theil's U statistic. And all these parameters are defined as following:

1) Mean absolute error:

$$\mathrm{M} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{7}$$

Where, At is the actual value and Ft is the forecast value
2) Root Mean Square error:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{8}$$

Where, is the mean and yi is the actual value
3) Theil's U Statistics

$$\mathrm{U} = \frac{\sqrt{\sum(A_i - F_t)^2}}{\sqrt{\sum A_i^2} - \sqrt{\sum F_t^2}} \tag{9}$$

Where, At is the actual value and Ft is the forecast value.

For higher order fuzzy time series we have performed with different orders to do the forecasting and did the comparison on basis of mean absolute percentage error of observed and predicted values, root mean square error and Theil's U statistic and found in tables attached above.

In table-1, we have done the comparison of different orders of Higher order Fuzzy time series model starting from order of 1 to 7. We have done the comparison on the basis of RMSE, MAPE and U value. For lower orders, the model does not perform well and when we increase the order the model performs better but if we go on increasing the order then overfitting of the model of takes place so to avoid overfitting and underfitting we need to choose the value of order in such a manner that the model performs well.

In table-2, we have shown the comparison of different fuzzy time series models which are Chen's model,ensemblerFTS model,Probabilistic model and higher order fuzzy model and found out that higher model performs best of all the models. In table-3, as we have forecasted the entire data by using higher order fuzzy time series so have shown the forecasted data of confirmed cases for the same.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have implemented several methods of Fuzzy Time series with unequal lengths of interval. In particular we have performed Higher Order fuzzy time series for forecasting COVID-19 confirmed cases, deaths, active cases and recovered cases in United states.And for comparison we have performed other fuzzy time series methods for forecasting of COVID-19 confirmed cases. In this paper we, in addition to fuzzy time series methods we have also performed traditional methods of forecasting which is Recurrent neural networks for comparison of results. From figure-5 we can conclude that in case of fuzzy time series higher order fuzzy time series model performs best and chen's model did not perform that well. In case of Higher fuzzy time series model the RMSE,MAPE and U all these errors are less in comparison to other models.From this we can conclude that higher order performs better than all other models as the predicted and forecasted values are almost same. Now, for traditional method which we can see in figure-6, Bi-directional LSTMs performs best among all three models as the we can see that the predicted value is almost equal to the forecasted values for the entire test dataset. Whereas Vanilla LSTMs model performed worst of all three model. From the above discussion we can conclude that by Bi-directional LSTMs performed best because in this case two LSTMs models were trained one from left to right and other one from right to left. Now, from this paper we found that the fuzzy time series models can also perform at par with the traditional forecasting methods and can produce good results or maybe better results if they some more amendments are done in the model. We have also came across to the fact that the model performs better in case of unequal interval lengths rather than equal because we can see from the nature of data that in the initial stage number of cases were very less whereas as the time increased there is exponential increase in number of cases so if we will take equal intervals then the results which we were getting they were not that good.

a) Future Work:: 1) For improvement of results, in case of defuzzification we can use Recurrent Neural networks with the help of which can help in achieving better results in case of forecasting.

2) In this paper, we have used only one factor for forecasting we can use different factors for forecasting which can be increase the accuracy.

3) We can try different partitioning techniques to get better performance of models for forecasting the data.

REFERENCES

[1] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – part I, Fuzzy Sets and Systems 54 (1993) 1–9.

[2] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – part II, Fuzzy Sets and Systems 62 (1994) 1–8.

[3] S.M. Chen, Forecasting enrollments based on fuzzy time series, Fuzzy Sets and Systems 81 (1996) 311–319.

[4] J.R. Hwang, S.M. Chen, C.H. Lee, Handling forecasting problems using fuzzy time series, Fuzzy Sets and Systems 100 (1998) 217–228.

[5] K. Huarng, Heuristic models of fuzzy time series for forecasting, Fuzzy Sets and Systems 123 (2001) 369–386.

[6] K. Huarng, Effective lengths of intervals to improve forecasting in fuzzy timeseries, Fuzzy Sets and Systems 123 (2001) 387–394.

[7] S.M. Chen, J.R. Hwang, Temperature prediction using fuzzy time series, IEEE Transactions on Systems, Man and Cybernetics 30 (2000) 263–275.

[8] K.-H. Huarng, T.H.-K. Yu, Y.W. Hsu, A multivariate heuristic model for fuzzy time-series forecasting, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 37 (2007) 836–846.

[9] C. Cheng, J. Chang, C. Yeh, Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost, Technological Forecasting and Social Change 73 (2006) 524–542.

[10] C.-C. Tsai, S.-J. Wu, Forecasting enrolments with high-order fuzzy time series, in: 19th International Conference of the North American, Fuzzy Information Processing Society, Atlanta, GA, 2000, pp. 196–200.

[11] S.-M. Chen, Forecasting enrollments based on high-order fuzzy time series, Cybernetics and Systems 33 (2002) 1–16.

[12] S.-M. Chen, N.-Y. Chung, Forecasting enrollments using high-order fuzzy time series and genetic algorithms, International Journal of Intelligent Systems 21 (2006) 485–501.

[13] T.-L. Chen, C.-H. Cheng, H.-J. Teoh, High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets, Physica A: Statistical Mechanics and its Applications 387 (2008) 876–888.

[14] S.-M. Chen, C.-D. Chen, Handling forecasting problems based on high-order fuzzy logical relationships, Expert Systems with Applications 38 (2011) 3857–3864.

[15] Zadeh, L. A., "The concept of a linguistic variable and its application to approximate reasoning-Part I", Information Sciences, Vol. 8, pp. 199-249, 1975.

[16] Silva, P. C. D. L., Alves, M. A., Severiano Jr, C. A., Vieira, G. L., Guimarães, F. G., Sadaei, H. J. (2017). Probabilistic forecasting with seasonal ensemble fuzzy time-series. In XIII Brazilian Congress on Computational Intelligence, Rio de Janeiro.

[17] Ahmed Mohammed, A., Aung, Z. (2016). Ensemble learning approach for probabilistic forecasting of solar power generation. Energies, 9(12), 1017.

[18] Boaisha, S. M., Amaitik, S. M. (2010). Forecasting model based on fuzzy time series approach. In Proceedings of the 10th International Arab Conference on Information Technology-ACIT (pp. 14-16).

[19] Song, Q., Chissom, B. S. (1993). Fuzzy time series and its models. Fuzzy sets and systems, 54(3), 269-277.

[20] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. Information sciences, 8(3), 199-249.

[21] Singh, P., Borah, B. (2013). High-order fuzzy-neuro expert system for time series forecasting. Knowledge-Based Systems, 46, 12-21.

[22] Severiano, C. A., Silva, P. C., Sadaei, H. J., Guimarães, F. G. (2017, July). Very short-term solar forecasting using fuzzy time series. In 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE) (pp. 1-6). IEEE.

[23]] Chen, S. M.—Chang, Y. C.: Multi-Variable Fuzzy Forecasting Based on Fuzzy Clustering and Fuzzy Rule Interpolation Techniques. Information Sciences, Vol. 180, 2010, No. 24, pp. 4772,4783

[24] H.S. Lee, M.T. Chou, Fuzzy forecasting based on fuzzy time series, International Journal of Computer Mathematics 81 (2004) 781–789.

[25] H.-K. Yu, Weighted fuzzy time series models for TAIEX forecasting, Physica A: Statistical Mechanics and its Applications 349 (2005) 609–624.

[26] C.H. Cheng, G.W. Cheng, J.W. Wang, Multi-attribute fuzzy time series method based on fuzzy clustering, Expert Systems with Applications 34 (2008) 1235–1242.

[27] C. Kai, F.F. Ping, C.W. Gang, A novel forecasting model of fuzzy time series based on k-means clustering, in: 2010 Second International Workshop on Education Technology and Computer Science, China, pp. 223–225.

[28] C.-H. Cheng, T.-L. Chen, L.-Y. Wei, A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting, Information Sciences 180 (2010) 1610–1629.

[29] I.-H. Kuo, S.-J. Horng, T.-W. Kao, T.-L. Lin, C.-L. Lee, Y. Pan, An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization, Expert Systems with Applications 36 (2010) 6108–6117.

[30] C. Aladag, M. Basaran, E. Egrioglu, U. Yolcu, V. Uslu, Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations, Expert Systems with Applications 36 (2009) 4228–4231.

[31] H. Teoh, C. Cheng, H. Chu, J. Chen, Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets, Data and Knowledge Engineering 67 (2008) 103–117.

[32] K. Bisht and S. Kumar, "Fuzzy time series forecasting method based on hesitant fuzzy sets," Expert Systems with Applications, vol. 64, pp. 557–568, 2016.

[33] K. K. Gupta and S. Kumar, "Hesitant probabilistic fuzzy set based time series forecasting method," Granular Computing, vol. 4, no. 4, pp. 739–758, 2019.

[34] K. Bisht and S. Kumar, "Hesitant fuzzy set based computational method for financial time series forecasting," Granular Computing, vol. 1–15, 2018.

[35] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

VII. Appendix

We have done experiments on additional data, as mentioned in the below tables these are the obtained by considering covid-19 Confirmed cases data of other 4 countries and comparing the metrics for each models with higher order FTS with order-6.

TABLE IV
Comparison of Fuzzy Time Series Model

Confirmed Cases of India				
Metrics	**Chen**	**EnsembleFTS**	**ProbabilisticWeightedFTS**	**Higher OrderFTS(Order-3)**
RMSE	0.37577801	0.205962835	0.034732624	0.035014346
MAPE	0.572814188	0.392266511	0.030980837	0.004599109
U	0.172735426	0.141266968	0.022850679	0.009457014

Confirmed Cases of Canada				
Metrics	**Chen**	**EnsembleFTS**	**ProbabilisticWeightedFTS**	**Higher OrderFTS(Order-3)**
RMSE	0.399008856	0.4023232	0.3453674	0. 3224256
MAPE	0.172735426	0.1743444	0.0534985	0.12457234
U	0.6889	0.5404	0.0795	0.0794

Confirmed Cases of Australia				
Metrics	**Chen**	**EnsembleFTS**	**ProbabilisticWeightedFTS**	**Higher OrderFTS(Order-3)**
RMSE	0.39875	0.32146	0.25784	0. 19698
MAPE	0.17915	0.14536	0.09479	0.05386
U	0.69657	0.53035	0.06924	0.05955

Confirmed Cases of China				
Metrics	**Chen**	**EnsembleFTS**	**ProbabilisticWeightedFTS**	**Higher OrderFTS(Order-3)**
RMSE	0.45898	0.40278	0.34591	0. 26573
MAPE	0.17279	0.12658	0.05392	0.02936
U	0.68225	0.57023	0.07621	0.03843

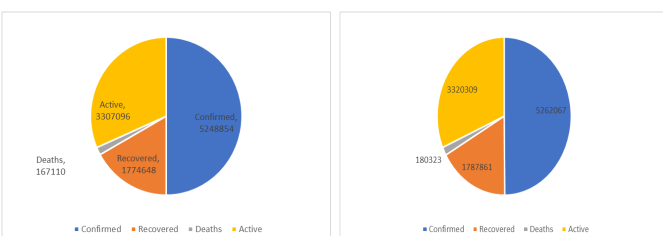


Fig. 6. (a) True values of US (b) Predicted values of US

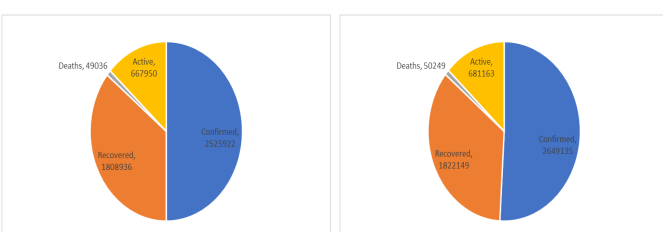


Fig. 7. (a) True values of India (b) Predicted values of India

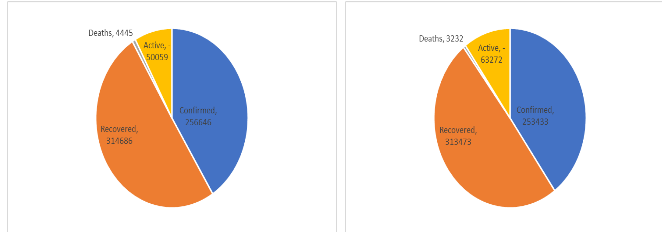


Fig. 8. (a) True values of Australia (b) Predicted values of Australia